Abstract

A method of estimating selectivity of a given string predicate in a database query. In the method selectivities of substrings of various substring lengths are estimated. For example, the selectivity of substrings between length 1 (or some constant q) to the length of the given string predicate may be estimated. The method then selects a candidate substring for each substring length b ased on e stimated selectivities of the substrings. The estimated selectivities of the candidate substrings are combined. The combined estimated selectivity of the candidate substrings is returned as the estimated selectivity of the given string predicate.